

Mismatch negativity at Fz in response to within-category changes of the vowel /i/

Ellen Marklund, Iris-Corinna Schwarz and Francisco Lacerda

The amplitude of the mismatch negativity response for acoustic within-category deviations in speech stimuli was investigated by presenting participants with different exemplars of the vowel /i/ in an odd-ball paradigm. The deviants differed from the standard either in terms of fundamental frequency, the first formant, or the second formant. Changes in fundamental frequency are generally more salient than changes in the first formant, which in turn are more salient than changes in the second formant. The mismatch negativity response was expected to reflect this with greater amplitude for more salient deviations. The fundamental frequency deviants did indeed result in greater amplitude than both first formant deviants and second formant deviants, but no difference was found between the first formant deviants and the second formant deviants. It is concluded that greater difference between standard and within-category deviants across different

acoustic dimensions results in greater mismatch negativity amplitude, suggesting that the processing of linguistically irrelevant changes in speech sounds may be processed similar to nonspeech sound changes. *NeuroReport* 25:756–759 © 2014 Wolters Kluwer Health | Lippincott Williams & Wilkins.

NeuroReport 2014, 25:756–759

Keywords: electroencephalography, mismatch negativity, phonetic processing, within-category deviants

Phonetics Laboratory, Department of Linguistics, Stockholm University, Stockholm, Sweden

Correspondence to Ellen Marklund, BA, Phonetics Laboratory, Department of Linguistics, Stockholm University, SE-106 91 Stockholm, Sweden
Tel: +46 8 6747347; fax: +46 8 155389; e-mail: ellen@ling.su.se

Received 12 November 2013 accepted 13 March 2014

Background

The auditory mismatch negativity (MMN) is an event-related potential component elicited in response to occasional changes in a sequence of otherwise similar auditory stimuli. It is calculated by subtracting the event-related potential from the frequently occurring stimuli (standards) from that of the rare differing stimuli (deviants), and can be seen as a negative shift in the resulting difference wave, typically peaking at about 150–250 ms after stimuli onset. The MMN component is found in central and frontocentral electrodes when using nose or mastoid electrodes as reference. The component is generated irrespective of the participants' level of attention to stimuli, and is therefore considered to represent an automatic change-detection process. Its generators are found in the frontal lobes and the supratemporal cortices [1].

The supratemporal component of the MMN is lateralized differently depending on the stimulus type, representing two different although sometimes parallel change-detection processes [2]. When stimuli are either nonspeech sounds or speech sounds with a linguistically irrelevant change between standard and deviant (e.g. between two exemplars of the same vowel with different fundamental frequency), change detection is of an acoustic nature and

bilaterally distributed, but when stimuli are speech sounds and the change between the standard and deviant is linguistically relevant (e.g. between two different vowels), a left-hemispheric phonemic change-detection process is also active [3–6]. Thus, it is not speech processing *per se* that causes the left-hemispheric activation, but detection of a change that is linguistically relevant, that is, the standard is perceived as one phoneme and the deviant as another.

The above-described processes are reflected in the peak amplitude and latency of the MMN response to different deviants. Although the amplitude and latency differences between deviants have been shown to result at least in part from the N1 component [7], this combination of MMN and N1 is commonly referred to as only MMN, and will be so also in the present manuscript. For nonspeech stimuli, the amplitude and latency vary systematically with the acoustic difference between standard and deviant stimulus, with greater difference resulting in greater amplitude and shorter latency [8,9]. When the stimuli, instead, are speech sounds, the amplitude reflects not only the acoustic difference between standard and deviant but also whether or not the phonemic change-detection process is activated. A deviant belonging to a different phonemic category than the standard results in a greater MMN amplitude than a deviant belonging to the same category as the standard sound, even when the deviants are acoustically equally different from the standard [10,11], suggesting that the linguistic processing contributes considerably toward

This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 3.0 License, where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially.

the overall amplitude of the response. The effect of the acoustic difference remains when stimuli are speech sounds as comparisons between different between-category deviants show the same pattern of greater acoustic difference resulting in greater amplitude of the MMN [12,13].

A number of studies have included within-category deviants with different magnitude of difference from the standard, but in most cases, no specific comparisons between these have been reported [14–16] because it was not relevant to the specific research questions. One exception is a recent study by Pakarinen *et al.* [14] in which larger magnitudes of difference in vowel pitch and vowel intensity were shown to correspond to greater MMN amplitude.

The aim of the present study is to extend previous findings on the effect of purely acoustic (i.e. phonetic) speech sound changes on the MMN amplitude, by comparing changes across different acoustic dimensions, thus creating high ecological validity. As the changes are not phonemic, it is expected that greater acoustic difference will result in greater MMN amplitude, even though the changes are not within a single acoustic dimension. The stimuli are different exemplars of the vowel /i/, with deviants of different acoustic distance from the standard. Deviants differ from the standard in different acoustic dimensions, either in terms of fundamental frequency (f0), the first formant (F1), or the second formant (F2). The hypothesis is that as f0-deviants in general are more salient than F1-deviants, which in turn are more salient than F2-deviants, the amplitude of the MMN will be higher for f0-deviants than for F1-deviants and higher for F1-deviants than for F2-deviants. In addition, the different deviant types will be presented in both single-deviant blocks and multiple-deviant blocks to test whether more salient deviants attenuate the response to less salient deviants.

Methods

Participants

The participants were 13 right-handed native speakers of Swedish (three women, mean age 29 years, range 25–39 years). They were given movie vouchers as compensation for their participation. One electroencephalography recording was excluded from the analysis because of technical failure during data collection. The study was approved by the Ethical Review Board, Karolinska Institutet (2011/955-31/1).

Stimuli

The stimuli consisted of serially synthesized four-formant versions of the vowel /i/, created in Praat 5.3.13 [17]. The vowels were 400 ms long with 50 ms fade in/out. The formant values were constant for the duration of the vowel, whereas f0 had a symmetrical linear rise/fall contour, peaking in the middle of the vowel at 110% of

the starting frequency and then back to the starting frequency. When f0 is denoted as 100 Hz, for example, it was 100 Hz at vowel onset, peaked at 110 Hz at 200 ms, and was back to 100 Hz at 400 ms (vowel offset). The values used for the first four formants of the standard stimulus were 255, 2190, 3150, and 3730 Hz [18], and f0 was set to 100 Hz at vowel onset and offset. The values of the deviant sounds differed from the standards in either f0, F1, or F2, whereas the third and fourth formant values were identical to those of the standard. For each deviant type (f0, F1, or F2), there were four versions, with f0-variations of ± 5 or ± 10 Hz, F1-variations of ± 25 or ± 50 Hz, or F2-variations of ± 100 or ± 200 Hz.

Experimental design

The experiment used an odd-ball design and consisted of five blocks, presented in a random order. Three of the blocks contained one type of deviant each (f0-, F1-, or F2-deviants), the fourth contained two types of deviants (F1- and F2-deviants), and the fifth contained deviants of all three types (f0-, F1-, and F2-deviants). At the beginning of each block, 10 standard stimuli were presented to establish them as standards. The following number of trials in each block depended on the number of deviant types presented; each deviant type was presented a total of 80 times (20 times per version) and the deviants always comprised 20% of the total number of trials (not including the initial standards). The stimuli were presented in a random order, except that two deviants were never presented successively. Stimulus onset asynchrony (onset-to-onset) was 1000 ms.

Procedure

During the experiment, participants were seated in front of a screen on which a muted film was playing. The stimulus sounds were presented through loudspeakers at a comfortable sound level and the participants were instructed not to pay attention to the sounds. Before the start of each block, electrode impedance was measured to ensure that it was less than 100 k Ω for each of the 128 electrodes. The total duration of the experiment session was ~ 2 h, including net application and impedance measurements. Because of technical problems, one of the participants listened to part of the F1-block twice, but only the second run was completed and included in the analysis.

Electroencephalography system

Data were collected and processed using NetStation 4.4 and high-impedance 128 electrode Hydrocel Sensor Nets (Electrical Geodesic Inc., Eugene, Oregon, USA). During recording, the signal was amplified using a Net Amps 300 amplifier (Electrical Geodesic Inc.) with a 20 kHz sampling rate, low-pass filtered at 4 kHz, and down-sampled to a 250 Hz sampling rate. The reference during recording was Cz.

Data processing

The data were filtered off-line using a band-pass filter of 1–40 Hz. Channels in which the voltage varied by more than 75 μV within a time window of –200 to 1000 ms around stimulus onset for more than half of the stimuli occurrences were marked as poor and then interpolated from surrounding channels [19]. Following this, the data were segmented into single trial epochs, 1000 ms long, from –100 to 900 ms relative to stimulus onset and divided into different categories for each deviant type by block type (a total of eight categories). All epochs in which the signal (averaged over 80 ms) exceeded $\pm 55 \mu\text{V}$ in 20 of the channels or more, or contained an eye-blink artifact ($\pm 140 \mu\text{V}$ in the horizontal electrooculogram channels) or an eye-movement artifact ($\pm 55 \mu\text{V}$ in the vertical electrooculogram channels) were marked as poor and excluded from further analysis. All participants had a minimum of 55% good epochs in each deviant-type category. The data were then rereferenced to the average of all channels and baseline corrected. All artifact-free pairs of any deviant and its immediately preceding standard were used in the analysis. The mean amplitude at electrode 11 (closely corresponding to Fz in the 10–20 system) during 150–250 ms after stimulus onset was calculated. Statistical tests were performed in SPSS 19 (International Business Machines Corp., Armonk, New York, USA).

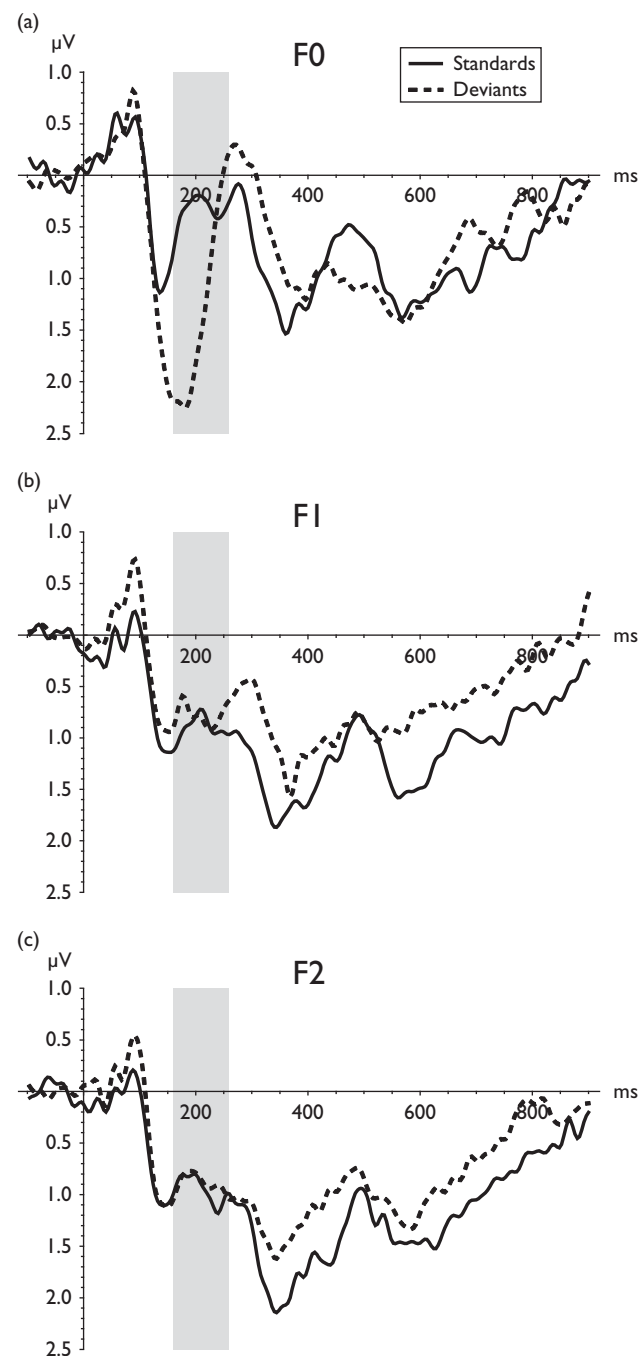
Results

A repeated-measures analysis of variance was performed on the amplitudes of deviants and standards, with the two within-participant factors deviant type (f0, F1, or F2) and block type (single, double, or triple). The difference between deviants and standards was highly significant [$F(1,5046) = 17.884$, $P < 0.001$], as was the interaction with deviant type [$F(2,5046) = 7.088$, $P = 0.001$]. There was neither interaction with block type [$F(2,5046) = 1.445$, $P = 0.236$] nor with block type and deviant type [$F(3,5046) = 2.188$, $P = 0.087$]. To investigate the effect of deviant type (Fig. 1), two-tailed paired-samples t -tests between standard and deviant were performed for each deviant type separately, showing a significant difference for f0-deviants [$t(1284) = -4.941$, $P < 0.001$] but not for F1- or F2-deviants [$t(1869) = -1.100$, $P = 0.272$, and $t(1898) = -0.935$, $P = 0.350$, respectively].

Discussion

The amplitude of the MMN response was investigated for within-category changes of different acoustic dimensions in speech stimuli. The hypothesis was that an acoustically larger difference would result in greater amplitude, as has been shown previously within a single acoustic dimension such as pitch or intensity [14]. The results of the present study are in line with this, with changes in fundamental frequency resulting in a significant MMN response for changes in fundamental frequency, but not for spectral changes. Thus, the same pattern is found for within-category changes that have

Fig. 1



The average ERP tracings at electrode 11 (Fz), pooled across blocks. For (a) f0-deviants, a difference between the standard response and deviant response can be seen, but not for (b) F1-deviants or (c) F2-deviants. Gray areas mark the time window used in the analysis (150–250 ms after stimulus onset). ERP, event-related potential.

previously been shown for between-category changes [11,12] and nonspeech sounds [8,9].

On the basis of the assumption that changes in F1 in general are more salient than changes in F2, it was hypothesized that the F1-deviants would elicit a greater

MMN than F2-deviants. However, no significant MMN response was found for either of these conditions. This could be a result of the relatively low number of trials and the variability of the deviants; there were four versions of each deviant type, with varying magnitude of difference to the standard. It is possible that larger F2-deviants were just as perceptually salient as smaller F1-deviants and that the average difference for all four versions combined was roughly equal for both types of spectral deviants. No effect on the MMN amplitude was found for the context (single-deviant or multiple-deviant blocks) in which deviants were presented.

Conclusion

The present study has shown that the relationship between the amplitude of the MMN and the magnitude of acoustic difference between the standard and the deviant is present for speech stimuli when the changes between standard and deviant are of different acoustic dimensions. Specifically, changes in fundamental frequency resulted in an MMN response, whereas changes in the first or the second formant did not. These findings, together with previous research, highlight the importance of differentiating between phonemic (linguistically relevant) and phonetic (linguistically irrelevant) processing of speech sounds: speech sounds are not necessarily processed differently from any other sound just because they are speech; processing differences primarily result instead from the presence of linguistically relevant information in the speech signal.

Acknowledgements

This study was funded by Stockholm University (135032).

Conflicts of interest

There are no conflicts of interest.

References

- Näätänen R, Paavilainen P, Rinne T, Alho K. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol* 2007; **118**:2544–2590.
- Näätänen R. The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology* 2001; **38**:1–21.
- Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huottilainen M, Iivonen A, et al. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 1997; **385**:432–434.
- Takegata R, Nakagawa S, Tonoike M, Näätänen R. Hemispheric processing of duration changes in speech and non-speech sounds. *Neuroreport* 2004; **15**:1683–1686.
- Kasai K, Yamada H, Kamio S, Nakagome K, Iwanami A, Fukuda M, et al. Brain lateralization for mismatch response to across- and within-category change of vowels. *Neuroreport* 2001; **12**:2467–2471.
- Shestakova A, Brattico E, Huottilainen M, Galunov V, Soloviev A, Sams M, et al. Abstract phoneme representations in the left temporal cortex: magnetic mismatch negativity study. *Neuroreport* 2002; **13**:1–4.
- Horváth J, Czigler I, Jacobsen T, Maess B, Schröger E, Winkler I. MMN or no MMN: no magnitude of deviance effect on the MMN amplitude. *Psychophysiology* 2008; **45**:60–69.
- Pakarinen S, Takegata R, Rinne T, Huottilainen M, Näätänen R. Measurement of extensive auditory discrimination profiles using the mismatch negativity (MMN) of the auditory event-related potential. *Clin Neurophysiol* 2007; **118**:177–185.
- Tiitinen H, May P, Reinikainen K, Näätänen R. Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature* 1994; **372**:90–92.
- Dehaene-Lambertz G. Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport* 1997; **8**:919–924.
- Sharma A, Dorman MF. Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *J Acoust Soc Am* 1999; **106**:1078–1083.
- Deguchi C, Chobert J, Brunellière A, Nguyen N, Colombo L, Besson M. Pre-attentive and attentive processing of French vowels. *Brain Res* 2010; **1366**:149–161.
- Sittiprapaporn W, Tervaniemi M, Chindaduangratn C, Kotchabhakdi N. Preattentive discrimination of across-category and within-category change in consonant vowel syllable. *Neuroreport* 2005; **16**:1513–1518.
- Pakarinen S, Teinonen T, Shestakova A, Kwon MS, Kujala T, Hämäläinen H, et al. Fast parametric evaluation of central speech-sound processing with mismatch negativity. *Int J Psychophysiol* 2013; **87**:103–110.
- Pakarinen S, Lovio R, Huottilainen M, Alku P, Näätänen R, Kujala T. Fast multi-feature paradigm for recording several mismatch negativities (MMNs) to phonetic and acoustic changes in speech sounds. *Biol Psychol* 2009; **82**:219–226.
- Partanen E, Vainio M, Kujala T, Huottilainen M. Linguistic multifeature MMN paradigm for extensive recording of auditory discrimination profiles. *Psychophysiology* 2011; **48**:1372–1380.
- Boersma P, Weenink D. Praat: doing phonetics by computer [Computer program], version 5.3.13. Available at: <http://www.praat.org/>. (Accessed 18 April 2012).
- Fant G, Henningsson G, Stålhammar U. Formant frequencies of Swedish vowels. *STL-QPSR* 1969; **10**:29–31.
- Perrin F, Pernier J, Bernard O, Giard MH, Echallier JF. Mapping of scalp potentials by surface spline interpolation. *Electroencephalogr Clin Neurophysiol* 1987; **66**:75–81.